1 Introduction

1.1 Attachment Trees

A *random attachment tree*, is a nested family $\{T_n\}_{n=1}^{\infty}$ of labeled, rooted trees obtained by assigning T_1 to be the tree with 1 vertex labelled v_1 and 0 edges and building $T_n + 1$ from T_n attaching vertex $v_n + 1$ by an edge to a randomly chosen vertex in T_n. A random recursive q-ary tree is random attachment tree where each vertex has maximum outdegree q. A random recursive tree (RRT) is a random attachment tree where each new vertex is attached *uniformly* at random to an existing vertex.

2 Attachment trees as a probability space

A probability space $(\Omega, \mathcal{F}, P)$ is a measure space where we say that Ω is a sample space, $\mathcal{F}$ is an event space (there exist events $E \in \mathcal{F}$), $P(E)$ is the probability of event E occuring and $P(\Omega) = 1$. In this section we will show that we can think of attachment trees as probability spaces.

2.1 A space of attachment trees

Clearly L is not the correct space for us to consider attachment trees. Another space we could consider is the infinite product space of $N_n := \{0, 1, 2, 3, \ldots, n\}$ for $n = 1, 2 \ldots \infty$ such that each $\mathbb{N}_n$ is equipt with the discrete topology.

$$\mathcal{L} = \prod_{i=1}^{\infty} \mathbb{N}_n. \tag{1}$$

We claim that any $l \in \mathcal{L}$ corrosponds to an attachment tree since such a tree is built iteratively by attaching vertex 2 to vertex 1 then attaching vertex 3 to vertex 2 with probability 0.5 or to vertex 1 with probability 0.5 etc. so that each $N \in \mathbb{N}_n$ represents the single vertex with a higher label that vertex $n+1$ is attached to. We can think of this as an infinite sequence of independen random variables.

Similarly the space of random recursive q-ary trees, $\mathcal{L}_q$ can be thought of as an infinite product space:

$$\mathcal{L}_q = \prod_{i=1}^{\infty} \Omega_n$$

such that $\Omega_n = \mathbb{N}_n$ for $n = 1, \ldots q - 1$ and $\Omega_n = \mathbb{N}_q$ subsequently.

2.2 A probability measure on attachment trees

Now that we have suitably defined the space of attachemnt trees, we also need to define a measure on this space so we appeal to a theorem of Tao [].

Theorem 1. *Let A be an arbitary set. For each $\alpha \in A$ let $(\Omega_\alpha, \mathcal{F}_\alpha, P_\alpha)$ be a probability space such that Ω_α is a locally compact, σ-compact metric space with Borel σ-algebra $\mathcal{F}_\alpha$ then there exists a unique probability measure*

$$P_A = \prod_{\alpha \in A} P_\alpha \text{ on } \left(\prod_{\alpha \in A} \Omega_\alpha, \prod_{\alpha \in A} \mathcal{F}_\alpha, \right)$$

Such that $P_A \left(\prod_{\alpha \in A} E_\alpha \right) = \prod_{\alpha \in A} P_\alpha(E_\alpha)$.

Furthermore, whenever $E_\alpha \in \mathcal{F}_\alpha$ one has $E_\alpha = \Omega_\alpha$ for all but finitely many α.

We aim to use Theorem 1 to prove that there exists a unique probability measure on the space $\mathcal{L}$ defined in section 2.1. Let $A = \mathbb{N}$ then for all $\alpha \in A$ we wish to prove the following Lemmas:

Lemma 2.1. *$(\mathbb{N}_\alpha, \mathcal{P}(\mathbb{N}_\alpha), \mu_\alpha)$ is a probability space where μ_α is the uniform probability measure and $\mathcal{P}(\mathbb{N}_\alpha)$ is the power set of $\mathbb{N}_\alpha$.*

Lemma 2.2. *$\mathbb{N}_\alpha$ is a locally compact space.*

Lemma 2.3. *$\mathbb{N}_\alpha$ is a σ-compact metric space.*

3 Pólya Urns

A generalised Pólya urn $\mathcal{U}$ contains a finite number of balls of finitely many possible types $1, 2, \ldots, q$. The content of the urn at time t is described by a vector $X_t = (X_{t,1}, X_{t,2}, \ldots X_{t,q})$ where each $X_{t,i}$ is the number of balls of type i in the urn at time t. We associate an *activity* a_i and a transition vector $\xi_i = (\xi_{i,1}, \xi_{i,2}, \ldots, \xi_{i,q})$ to each type. A generalised Pólya urn is an evolving process with initial content X_0 and at subsequent times a ball is drawn from the urn. We also assume that $\mathbb{E}|X_0|^2$. At time t a ball of type i is drawn with probability

$$P_{i,t} = \frac{a_i X_{t-1,i}}{\sum_{j=1}^q X_{t-1,j}}$$

The drawn ball is then returned to the urn with ξ_{ij} balls of type j for $j = 1, 2, \ldots, q$. In general we also require the following further conditions:

$$\xi_{i,j} \geq 0 \text{ if } j \neq i \tag{2}$$

$$\xi_{i,i} \geq -1 \tag{3}$$

Assume that a ball of type i is chosen from an urn, Equation 2 ensures that no balls of a type *other* than i is removed from the urn and Equation 3 means that at most one ball of type i is removed from the urn. Together these conditions prevent teh removal of a ball that does not exist from a generalised Pólya urn.

Remark 1. *If $a_i = 1$ for every type i then at each time a ball is drawn from the urn it has a uniformly random type.*

3.1 Properties of Pólya urns

The crux of this section is Theorem 2 which is a limit theorem for Pólya urns. In order to state this theorem we require some further concepts from urn theory.

To every Pólya urn we can associate a matrix A such that the j^{th} column of A is defined to be $a_j\mathbb{E}(\xi_j)$ where $\mathbb{E}$ denotes expectation.

We will write $i \succ j$ if it is possible to find a ball of type j in an urn beginning with a single ball of type i. We say that a type i is *dominating* if $i \succ j$ for every type $j = 1, 2, \dots q$. Note that $\succ$ is a transitive and reflexive relation so it partitions the set of types in to equivilence classes $C_1, C_2, \dots, C_r$ where types $i, j \in C_k$ if $i \succ j$ and $j \succ i$. We say that some class C_k is dominating if some $i \in C_k$ is dominating.

We say that a generalised Pólya urn becomes *essentially extinct* if at some time t there does not exist a ball of dominating type.

Theorem 2. *Let A be the matrix associated to some non-essentially extinct Pólya urn process such that:*

(A1) Equations 2 and 3 are satisfied.

(A2) $\mathbb{E}(\xi_{ij}) < \infty$ for all $i, j = 1, 2, \dots, q$.

(A3) There exists a largest real eigenvalue λ_1 of A which is positive.

(A4) λ_1 is simple.

(A5) There exists a dominating type, i, and $X_{0,i} > 0$.

(A6) λ_1 belongs to dominating type.

Let v_1 be the right eigenvector associated with λ_1 then:

$$n^{-1}X_n \to \lambda_1 v_1 \text{ almost surely as } n \to \infty$$

Janson uses Theorem 2 to prove that if $X_{n,i}$ is the number of vertices with outdegree $i (\geq 0)$ in a RRT on n vertices then:

Theorem 3. *In the limit as $n \to \infty$, $\frac{X_{n,i}}{n} \to 2^{-i-1}$ almost surely.*

Theorem 3 is an example of the strong law for large numbers. Loosely speaking, given any infinite RRT, we can think of $\frac{X_{n,i}}{n}$ as a trajectory which gets ever closer to the expected outdegree of 2^{-i-1}.

4 A specific Pólya urn

??

In this section we will describe a generalized Pólya urn , $\mathcal{U}_m$ that describes the distribution of random recurssive d-ary tree motifs. Subsequently we will hit $\mathcal{U}_m$ with Theorem 2 to prove a limiting thoerem for network motifs.

Let $\{T_n\}_{n=1}^{\infty}$ be a random recursive d-ary tree process. At time n any vertex $v \in V(T_n)$ is incident to l leaves where $0 \leq l \leq d$. We call the induced subtree of T_n with a hubnode adjacent to l leaves such that each leaf has higher label than the hub an l-star (an l-star is an example of a network motif). By taking l to be maximal partition $V(T_n)$ can be partitioned into $d+1$ sets of vertices contained in induced l-stars for $l = 1, 2, 3, d$ and the set of vertices not contained in any l-star.

Balls in $\mathcal{U}_m$ may take one of $d+1$ types that correspond to the aforementioned partition. In particular types $2, 3, \ldots d+1$ correspond to vertices in $(d-1)$-stars and type 1 corresponds to the remaining vertices (for conveniance we will refer to these vertices as 0-stars). Since an l-star contains $l+1$ vertices we set activities $a_i = i$ for $i = 1, 2, \ldots d$ and $a_{d+1} = d$ for the reasons we give below.

At time $n+1$ vertex v_{n+1} is attached to i is attached via an edge to a vertex $v \in V(T_n)$. In order that the distribution of network motifs is the same as the distribution of types the probability that v is contained in an induces $i+1$-star must be the same as the probability that a ball of type i is drawn from $\mathcal{U}_m$ at time $n+1$. Furthermore, we claim that the appropriate transition vectors are as follows:

$$
\begin{aligned}
\xi_1 &= (-1, 1, 0, \ldots, 0) \\
\xi_2 &= (\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, 0, \ldots, 0) \\
\xi_3 &= (-0, \frac{4}{3}, -1, \frac{1}{3}, 0, \ldots, 0) \\
\xi_i &= (0, \frac{i-1}{i}, 0, \ldots, 0, \frac{i-1}{i}, -1, \frac{1}{I}, O, \ldots, 0) \\
\xi_{d+1} &= (0, 1, 0, \ldots, 0, 1, -1)
\end{aligned}
$$

of claim. For $i = 2$ assume that T_n and X_n are as usual and that one draws a ball of type 2 from U_n. Then equivilently at time $n+1$ one attaches vertex v_{n+1} via an edge to an induced 1-star of T_n for clarity let that 1-star have hub h and leaf l. With probability $\frac{1}{2}$ vertex v_{n+1} is attached to h, in which case a 2-star is created. Similarly with probability $\frac{1}{2}$ veretx v_{n+1} is attached to l in which case a 1-star and a 0-star are created. See Diagram for further details. Vectors ξ_i are built in an analogous way for $i = 3, 4, \ldots, d$. If a ball of type $d+1$ is drawn from $\mathcal{U}_m$ we could equivilently imagine that a vertex is connected via an edge to a d-star. Since each T_n is a d-ary tree $v_n + 1$ *must* be attached to a leaf of that d-star (this is why we set $a_d + 1 = d$. $\square$

5 Further Background

In order to prove Theorem 4 we require further background from linear algebra such as the Gershgorin Circle Theorem (Theorem 5.1) and a short introduction to Perron-Frobenius theory.

5.1 The Gershgorin circle Theorem

Theorem 5.1. *Gershgorin circle theorem Let $A \in M_n(\mathbb{C}$, and define*

$$R_i = \sum_{i=1, i \neq j}^{n} |a_{ij}|$$

Then each eigenvalue of A is in at least one of the disks

$$D_i = \{z : |z - a_{ii}| \leq R_i$$

Somewhat suprisingly the Gershgorin circle theorem gives us a bound on the values of the eigenvalues; informally it tells us that the eigenvalues cannot be too far from the diagonal elements of A.

5.2 Nonnegative matrices

The reference for this section is [**?**]. A square matrix $A = \{a_{ij}\}$ is said to be *nonnegative* if every element $a_{ij} \geq 0$ and we write $A \geq 0$. We say that a nonnegative matrix A is *irreducible* if for every a_{ij} there exists some $n \in \mathbb{N}$ such that $(a_{ij})^n > 0$.

A square matrix $A = \{a_{ij}\}$ is said to be Metzler if for all $i \neq j$, $a_{ij} \geq 0$. A Metzler matrix $A \in M_n(\mathbb{C})$ is related to a nonnegative matrix T by:

$$T = \mu I + A$$

for some large enough $\mu \in \mathbb{R}$.

Definition 5.2. A Metzler matrix A is is said to be irreducible if the related nonnegative matrix T is irreducible.

The theory of nonnegative matrices has been widely studied and Perrron-Frobenius theory yields the following theorem.

Theorem 5.3. *Let A be an irreducible Metzler square matrix.*

- *(i) Matrix A has an eigenvalue λ_1 such that $\lambda_1 \in \mathbb{R}$.*
- *(ii) λ_1 is associated with strictly positive left and right eigenvectors.*
- *(iii) $\lambda_1 > Re(\lambda)$ for any other other eigenvalue λ of A.*
- *(iv) λ_1 is simple.*

Theorem 5.4. *[**?**] Let $A = \{a_i j\}$ be an $n \times n$ matrix. Associate to A a directed graph G_A on n vertices with a directed edge from i to j whenever $a_{ij} > 0$. Then A is irreducible if and only if G_A is strongly connected.*

Theorem 4. *If $A_{d+1} = \{a_i j\}$ be the $(d+1) \times (d+1)$ matrix associated with the generalized Pólya urn U_m then A_{d+1} satisfies (A1)-(A6).*

Proof. (A1) True by construction.

(A2) true since d is finite $\xi_{ij} \leq d + 1 < \infty$.

(A3) Clearly 1 is an eigenvalue of A_{d+1} with left eigenvector $u_1 = (1, 1, \ldots, 1)$. We claim that 1 is the largest real eigenvector of A_q and to prove this claim we appeal to the Gershgorin Circle Theorem. Note that for each column in A_q, the sum of the diagonal elements:

$$C_1 = 2C_2 \qquad = 2C_i = \sum_{j \neq i} |a_{ij}| = i + 1 \text{ if } 3 \leq i \leq d + 1$$

Let $D_i = D(a_{ii}, C_i)$ be the disk in $\mathbb{C}$ centred at a_{ii} with radius C_i. Note that $D_1 = D(-1, 2)$, $D_2 = D(-1, 2)$ and $D_i = (-i, i + 1)$ for $i = 3, 4, 5, \ldots, d + 1$. Let $\mathcal{D} = \bigcup_i D_i(a_{ii}, C_i)$ (see Figure **??** for an image of $\mathcal{D}$). The largest real number in D is clearly 1, therefore by the Gergorin Circle Theorem the largest real eigenvalue λ_1 of A_{d+1} is 1.

(A4) Since no off diagonal element of A_{d+1} is negative A_{d+1} is a Metzler matrix. Therefore A_{d+1} can be associated with a nonnegative matrix T by writing:

$$T_\mu = \mu I + A_{d+1}$$

For some choice of $\mu \in \mathbb{R}$. We make the choice $\mu = d + 1$ so that $T_{d=1}$ is a nonnegative matrix. By Theorem 5.2 A_{d+1} is irreducible if T_{d+1} is irreducible.

Note that all the subdiagonal and superdiagonal entries of $T_d + 1$ are positive so the graph G_A built in the way described in Theorem 5.4 is strongly connected so A_{d+1} is irreducible. By Theorem 5.3 1 is a simple eigenvalue.

(A5) True since A_{d+1} is irreducible and $X_0 \neq 0$.

(A6) True since A_{d+1} is irreducible and $X_0 \neq 0$.

□